

DISCOVERY
OF
DIGITAL INFORMATION

Ronald J. Hedges, U.S.M.J.
Martin Luther King, Jr., Federal
Building and Courthouse
50 Walnut Street
Newark, New Jersey 07101
Judge_Ronald_Hedges@njd.uscourts.gov

TABLE OF CONTENTS

I.	WHAT DOES “DIGITAL INFORMATION” ENCOMPASS?.....	1
II.	WHEN TO BEGIN TO “THINK DIGITAL”.....	7
III.	DIGITAL INFORMATION AND RULE 26(f).....	9
IV.	DIGITAL INFORMATION AND RULE 26(a)(1).....	19
V.	DISCOVERY.....	21
VI.	COST-BEARING: THREE APPROACHES.....	28
VII.	AVOIDING PROBLEMS: SOME SUGGESTIONS	44
VIII.	CONCLUSION.....	51.

DISCOVERY OF DIGITAL INFORMATION¹

I. WHAT DOES “DIGITAL INFORMATION” ENCOMPASS?

A. Is digital information different?

“Computer files, including e-mails, are discoverable. *** . However, the Court is not persuaded by the plaintiffs’ attempt to equate traditional paper-based discovery with the discovery of e-mail files . Several commentators have noted important differences between the two. ***. Chief among these differences is the sheer volume of electronic information. E-mails have replaced other forms of communication besides just paper-based communication. Many informal messages that were previously relayed by telephone or at the water cooler are now sent via e-mail. Additionally, computers have the ability to capture several copies (or drafts) of the same e-mail, thus multiplying the volume of documents. All of these e-mails must be scanned for both relevance and privilege. Also, unlike most paper-based discovery, archived e-mails typically lack a coherent filing system. Moreover, dated archival systems commonly store information on magnetic tapes which have become obsolete. Thus, parties incur additional costs in translating the data from the tapes into useable form. One commentator has suggested that given the extraordinary costs of converting obsolete backup tapes into useable form, the requesting party should be required to show

¹For useful sources of information on digital discovery, see Digital Discovery & e-Evidence, a monthly publication of Pike & Fischer, Inc., and the unofficial web site created by Ken Withers of the Federal Judicial Center at <http://www.kenwithers.com>. See also, for a helpful “primer” on various aspects of electronic information, the two-part series of articles in the July and August, 2002 issues of The Federal Lawyer.

that production will likely result in the discovery of relevant information.” Byers v. Illinois State Police, 2002 U.S. Dist LEXIS 9861, *31-33 (N.D. Ill. May 31) (citations omitted).

B. On March 19, 2003, the Sedona Conference published “The Sedona Principles: Best Practices Recommendations and Principles for Addressing Electronic Document Production.”² The Sedona Principles included definitions, as described in Vol. 3, No. 4, Digital Discovery & e-Evidence 10 (April, 2003):

“Understanding technical terms is the first hurdle to overcome in mastering electronic evidence. To that end, the Sedona Principles are accompanied by a glossary of words and phrases. Here are the Sedona definitions of some of the less familiar terms.

Distributed Data: Distributed Data is that information belonging to an organization which resides on portable media and non-local devices such as home computers, laptop computers, floppy disks, CD-ROMS, personal digital assistants (‘PDAs’), wireless communication devices (e.g.,

²The Sedona Principles are available at <http://www.thesedonaconference.org>.

The American Bar Association has also been active in the area of electronic discovery. In 1999, its House of Delegates adopted “Civil Discovery Standards,” two of which addressed electronic discovery (available at <http://www.abanet.org/litigation/taskforces/standards.html>). On November 17, 2003, the Task Force on Electronic Discovery of the ABA’s Section of Litigation released for public comment new proposed standards that “focus on privilege problems and making discovery conferences and meet-and-confer effective.” G. Joseph, “Electronic Discovery Standards,” National L.J. 30 (Nov. 24, 2003). The proposed standards are available at <http://www.abanet.org/litigation/taskforces/electronic/home.html>.

Blackberry), zip drives, Internet repositories such as e-mail hosted by Internet service providers or portals, web pages, and the like. Distributed data also includes data held by third parties such as application service providers and business partners.

Forensic Copy: A Forensic Copy is an exact bit-by-bit copy of the entire physical hard drive of a computer system, including slack and unallocated space.

Legacy Data: Legacy Data is information the development of which an organization may have invested significant resources to and that has retained its importance, but has been created or stored by the use of software and/or hardware that has been rendered outmoded or obsolete.

Metadata: Metadata is information about a particular data set which describes how, when and by whom it was collected, created, accessed, and modified and how it is formatted. Some metadata, such as file dates and sizes, can easily be seen by users; other metadata can be hidden or embedded and is unavailable to computer users who are not technically adept. Metadata is generally not reproduced in full form when a document is printed. (Typically referred to by the not highly informative ‘shorthand’ phrase ‘data about data,’ describing the content, quality, condition, history, and other characteristics of the data.)

Residual Data: Residual Data (sometimes referred to as ‘Ambient Data’) refers to data that is not active on a computer system. Residual data includes (1) data found on media free space; (2) data found in the file slack space; and (3) data within files that have functionally been deleted in that it is not visible using the application with which the file was created, without use of undelete or special data recovery techniques.

Migrated Data: Migrated data is information that has been moved from one database or format to another, usually as a result of a change from one hardware or software technology to another.”

C. The Manual for Complex Litigation (Fourth Edition)³ assumes that, “[f]or the most part,” digital or electronic information will be “generated and maintained in the ordinary course of business.” However,

“Other data are generated and stored as a byproduct of the various information technologies commonly employed by parties in the ordinary course of business, but not routinely retrieved and used for business purposes. Such data include the following:

.Metadata, or ‘information about information.’ ***

.System data, or information generated and maintained by the computer itself. The computer records a variety of routine transactions and functions, including password access requests, the creation or deletion of files and directories, maintenance functions, and access to and from other computers, printers, or communication devices.

.Backup data, generally store offline on tapes or disks. Backup data are created and maintained for short-term disaster recovery, not for retrieving particular files, databases, or programs. These tapes or disks must be restored to the system from which they were recorded, or to a similar hardware and software environment, before any data can be accessed.

.Files purposely deleted by a computer user. Deleted files are seldom actually deleted from the computer hard drive. The operating system renames and marks them for eventual overwriting, should that particular space on the computer hard drive be needed. The files are recoverable only with expert intervention.

.Residual data that exist in bits and pieces throughout a computer hard drive. Analogous to the data on crumpled

³Hereinafter “Manual.” To be published in 2004.

newspapers used to pack shipping boxes, these data are also recoverable with expert intervention.

Each of these categories of computer data may contain information within the scope of discovery. The above categories are listed by order of potential relevance and in ascending order of cost and burden to recover and produce.” Manual, §11.446.

D. In Zubulake v. UBS Warburg LLC, 217 F.R.D. 309 (S.D.N.Y. 2003), the court discussed electronic data storage media. Here are the descriptions of those media, as summarized in Vol. 3, No. 6, Digital Discovery & e-Evidence 6 (June, 2003):

“Here are the full descriptions of electronic data storage media, taken from the Zubulake decision. The listings are in order from most to least accessible; citations have been omitted.

.Active, online data: Online storage is generally provided by magnetic disk. It is used in the very active stages of an electronic record’s life—when it is being created or received and processed, as well as when the access frequency is high and the required speed of access is very fast, i.e., milliseconds. Examples of online data include hard drives.

.Nearline data: This typically consists of a robotic storage device (robotic library) that houses removable media, uses robotic arms to access the media, and uses multiple read/write devices to store and retrieve records. Access speeds can range from as low as milliseconds if the media is already in a read device, up to 10-30 seconds for optical disk technology, and between 20-120 seconds for sequentially searched media, such as magnetic tape. Examples include optical disks.

.Offline storage/archives: This is removable optical disk or magnetic tape media, which can be labeled and stored in a shelf or rack. Offline storage of electronic records is traditionally used for making disaster copies of records and also for records considered ‘archival’ in that their likelihood of retrieval is minimal. Accessibility to offline media involves manual intervention and is much slower than online or nearline storage. Access speed may be minutes, hours or

even days, depending on the access–effectiveness of the storage facility. The principle difference between nearline data and offline data is that offline data lacks ‘the coordinated control of an intelligent disk subsystem,’ and is, in the lingo, JBOD (‘Just a Bunch Of Disks’).

.Backup tapes: A device, like a tape recorder, that reads data from and writes it onto a tape. Tape drives have data capacities of anywhere from a few hundred kilobytes to several gigabytes. Their transfer speeds also vary considerably. ... The disadvantage of tape drives is that they are sequential-access devices, which means that to read any particular block of data, you need to read all the preceding blocks. As a result, [t]he data on a backup tape are not organized for retrieval of individual documents or files [because] ... the organization of the data mirrors the computer’s structure, not the human records management structure. Backup tapes also typically employ some sort of data compression, permitting more data to be stored on each tape, but also making restoration more time-consuming and expensive, especially given the lack of uniform standard governing data compression.

.Erased, fragmented or damaged data: When a file is first created and saved, it is laid down on the [storage media] in contiguous clusters. ... As files are erased, their clusters are made available again as free space. Eventually, some newly created files become larger than the remaining contiguous free space. These files are then broken up and randomly placed throughout the disk. Such broken-up files are said to be ‘fragmented,’ and along with damaged and erased data can only be accessed after significant processing.”

E. Conclusion? “The complexity and rapidly changing character of technology for the management of computerized materials may make it appropriate for the judge to seek the assistance of a special master or neutral expert, or call on the parties to provide the judge with expert assistance, in the form of briefings on the relevant technological issues.” Manual, §11.446.

II. WHEN TO BEGIN TO “THINK DIGITAL”

A. Rule 11(a) requires that, “[e]very pleading, written motion, and other paper shall be signed by at least one attorney of record ***.”

B. Rule 11(b) provides that, “[b]y presenting to the court *** a pleading, written motion, or other paper, an attorney *** is certifying that, to the best of the person’s knowledge, information, and belief, formed after an inquiry reasonable under the circumstances,--

(1) it is not being presented for any improper purpose, such as to harass or to cause unnecessary delay or needless increase in the cost of litigation;

(2) the claims, defenses, and other legal contentions therein are warranted by existing law or by a nonfrivolous argument for the extension, modification, or reversal of existing law or the establishment of new law;

(3) the allegations and other factual contentions have evidentiary support or, if specifically so identified, are likely to have evidentiary support after a reasonable opportunity for further investigation or discovery; and

(4) the denials of factual contentions are warranted on the evidence or, if specifically so identified, are reasonably based on a lack of information or belief.”

C. The language of Rule 11 “stresses the need for some prefiling inquiry into both the facts and the law to satisfy the affirmative duty imposed by the rule. The standard is one of reasonableness under the circumstances.” Advisory Committee Note to 1983 amendment to Rule 11. Rule 11 “continues to require litigants to ‘stop-and-think’ before initially making legal or factual contentions.” Advisory Committee Note to 1993 amendment to Rule 11(b) and (c).

D. Why is knowledge of information in electronic format needed at earliest stage of litigation?

1. Ensure that there is an “off switch” for any deletion of data.
2. Comply with Rule 11.⁴

3. Prepare for Rule 26(f) conference.
4. Prepare for Rule 26(a)(1) disclosures.

III. DIGITAL INFORMATION AND RULE 26(f)

A. Why should discovery of electronic information be considered as early as possible? Here is what the Manual says:

“Computerized data have become commonplace in litigation. The sheer volume of such data, when compared to conventional paper documentation, can be staggering. A floppy disk, with 1.44 megabytes, is the equivalent of 720 typewritten pages of plain text. A CD-ROM, with 650 megabytes, can hold up to 325,000 typewritten pages. One gigabyte is the equivalent of 500,000 typewritten pages. Large corporate computer networks create backup data measure in terabytes, or 1,000,000 megabytes; each terabyte represents the equivalent of 500 billion typewritten pages of plain text.

Digital or electronic information can be stored in any of the following: mainframe computers, network servers, personal computers, hand-held devices, automobiles, or household appliances; or it can be accessible via the Internet, from private networks, or from third parties. Any discovery plan must address issues relating to such information, including the search for it and its location, retrieval, form of production, inspection, preservation, and use at trial.”

* * *

“There are several reasons to encourage parties to produce and exchange data in electronic form:

.discovery requests may themselves be transmitted in computer-accessible form—interrogatories served on computer disks, for example, could then be answered using the same disk, avoiding the need to retype them;

.production of computer data on disks, CD-ROMs, or by file transfers significantly reduces the costs of copying, transport, storage, and management—protocols may be established by the parties to facilitate the

handling of documents from initial production to use in depositions and pretrial procedures to presentation at trial;

.computerized data are far more easily searched, located, and organized than paper data; and

.computerized data may form the contents for a common document depository (see section 11.444).

The goal is to maximize these potential advantages while minimizing the potential problems of incompatibility among various computer systems, programs, and data, and minimizing problems with intrusiveness, data integrity, and information overload.” Manual, §11.446.

B. Rule 26(f) requires the parties to confer:

“(f) Conference of Parties; Planning for Discovery. Except in categories of proceedings exempted from initial disclosure under Rule 26(a)(1)(E) or when otherwise ordered, the parties must, as soon as practicable and in any event at least 21 days before a scheduling conference is held or a scheduling order is due under Rule 16(b), confer to consider the nature and basis of their claims and defenses and the possibilities for a prompt settlement or resolution of the case, to make or arrange for the disclosures required by Rule 26(a)(1), and to develop a proposed discovery plan that indicates the parties’ views and proposals concerning:

(1) what changes should be made in the timing, form, or requirement for disclosures under Rule 26(a), including a statement as to when disclosures under Rule 26(a)(1) were made or will be made;

(2) the subjects on which discovery may be needed, when discovery should be completed, and whether discovery should be conducted in phases or be limited to or focused upon particular issues;

(3) what changes should be made in the limitations on discovery imposed under these rules or by local rule,

and what other limitations should be imposed; and

(4) any other orders that should be entered by the court under Rule 26(c) or under Rule 16(b) and (c).

The attorneys of record and all unrepresented parties that have appeared in the case are jointly responsible for arranging the conference, for attempting in good faith to agree on the proposed discovery plan, and for submitting to the court within 14 days after the conference a written report outlining the plan. A court may order that the parties or attorneys attend the conference in person. If necessary to comply with its expedited schedule for Rule 16(b) conferences, a court may by local rule (I) require that the conference between the parties occur fewer than 21 days before the scheduling conference is held or a scheduling order is due under Rule 16(b), and (ii) require that the written report outlining the discovery plan be filed fewer than 14 days after the conference between the parties, or excuse the parties from submitting a written report and permit them to report orally on their discovery plan at the Rule 16(b) conference.”

C. The Rule 26(f) conference is the first opportunity to discuss electronic information with adversaries. Some district courts require the subject to be addressed:

1. Eastern and Western Districts of Arkansas Local Civil Rule 26.1:

“The Fed. R. Civ. P. 26(f) report filed with the court must contain the parties’ views and proposals regarding the following:

* * *

4. Whether any party will likely be requested to disclose or produce information from electronic or computer-based media. If so:

a. whether disclosure or production will be limited to data reasonably available to the parties in the ordinary course of business;

b. the anticipated scope, cost and time required for disclosure or production of data beyond what is reasonably available to the parties in the ordinary course of business;

c. the format and media agreed to by the parties for the production of such data as well as agreed procedures or such production;

d. whether reasonable measures have been taken to preserve potentially discoverable data from alteration or destruction in the ordinary course of business or otherwise;

e. other problems which the parties anticipate may arise in connection with electronic or computer-based discovery.”

2. District of New Jersey Local Civil Rule 26.1(b)(2):

“The parties shall submit their Fed. R. Civ. P. 26(f) discovery plan containing the parties’ views and proposals regarding the following:

(d) whether any party will likely request or produce computer-based or other digital information, and if so, the parties’ discussions of the issues listed under the Duty to Meet and Confer in L. Civ. R. 26.1(d)(3) below ***.”

3. District of New Jersey Local Civil Rule 26.1(d):

“(1) Duty to Investigate and Disclose. Prior to a Fed. R. Civ. P. 26(f) conference, counsel shall review with the client the client’s information management systems including computer-based and other digital systems, in order to understand how information is stored and how it can be retrieved. To determine what must be disclosed pursuant to Fed. R. Civ. P. 26(a) (1), counsel shall further review with the client the client’s information files, including currently maintained computer files as well as historical, archival, back-up, and legacy computer files, whether in current or historic media or formats, such as digital evidence which may be used

to support claims or defenses. Counsel shall also identify a person or persons with knowledge about the client's information management systems, including computer-based and other digital systems, with the ability to facilitate, through counsel, reasonably anticipated discovery.

(2) Duty to Notify. A party seeking discovery of computer-based or other digital information shall notify the opposing party as soon as possible, but no later than the Fed. R. Civ. P. 26(f) conference, and identify as clearly as possible the categories of information which may be sought. A party may supplement its request for computer-based and other digital information as soon as possible upon receipt of new information relating to digital evidence.

(3) Duty to Meet and Confer. During the Fed. R. Civ. P. 26(f) conference, the parties shall confer and attempt to agree on computer-based and other digital discovery matters, including the following:

(a) Preservation and production of digital information; procedures to deal with inadvertent production of privileged information; whether restoration of deleted digital information may be necessary; whether back up or historic legacy data is within the scope of discovery and the media, format, and procedures for producing digital information;

(b) Who will bear the costs of preservation, production, and restoration (if necessary) of any digital discovery.”

4. District of Wyoming Local Civil Rule 26.1(d)(3):

“(A) Duty to Notify. A party seeking discovery of computer-based information shall notify the opposing party immediately, but no later than the Fed. R. Civ. P. 26(f) conference of that fact and identify as clearly as possible the categories of information which may be sought.

(B) Duty to Meet and Confer. The parties shall meet and confer regarding the following matters during the Fed. R. Civ. P. 26(f) conference;

(I) Computer-based information (in general). Counsel shall attempt to agree on steps the parties will take to segregate and preserve computer-based information in order to avoid accusations of spoliation;

(ii) E-mail information. Counsel shall attempt to agree as to the scope of e-mail discovery and attempt to agree upon an e-mail search protocol. This should include an agreement regarding inadvertent production of privilege e-mail messages.

(iii) Deleted information. Counsel shall confer and attempt to agree whether or not restoration of deleted information may be necessary, the extent to which restoration of deleted information is needed, and who will bear the costs of restoration; and

(iv) Back-up data. Counsel shall attempt to agree whether or not back-up data may be necessary, the extent to which back-up data is needed and who will bear the cost of obtaining back-up data.”

D. Rule 26(f) affords the opportunity to, among other things:

1. Inquire into what information adversaries have in electronic format and how expensive production will be.⁵

2. Inquire into who is most knowledgeable about an adversaries’ electronic information systems.

3. Discuss preservation of electronic data.⁶ What the Manual says about preservation orders:

“Before discovery starts, and perhaps before the initial conference, consider whether to enter an order requiring the parties to preserve and retain documents, files, data, and records that may be relevant to the litigation. Because such an order may interfere with the normal operations of the

parties and impose unforeseen burdens, it is advisable to discuss with counsel at the first opportunity about the need for a preservation order and, if one is needed, the scope, duration, method of data preservation, and other terms that will best preserve relevant matter without imposing undue burdens.

A blanket preservation order may be prohibitively expensive and unduly burdensome for parties dependent on computer systems for their day-to-day operations. In addition, a preservation order will likely be ineffective if it is formulated without reliable information from the responding party regarding what data-management systems are already in place, the volume of data affected, and the costs and technical feasibility of implementation. The following are among the points to consider in formulating an effective data-preservation order:

.Continued operation of computers and computer networks in the routine course of business may alter or destroy existing data, but a data preservation order prohibiting operation of the computers absolutely would effectively shut down the responding party's business operations. Such an order requires the parties to define the scope of contemplated discovery as narrowly as possible, identify the particular computers or network servers affected, and agree on a method for data preservation, such as creating an image of the hard drive or duplicating particular data on removable media, thereby minimizing cost and intrusiveness and the downtime of the computers involved.

.Routine system backups for disaster recovery purposes may incidentally preserve data subject to discovery, but recovery of relevant data from nonarchival backups is costly and inefficient, and a data-preservation order that requires the accumulation of such backups beyond their usual short retention period may needlessly increase the scope and cost of discovery. An order for the preservation of backup data obliges the parties to define the scope of contemplated discovery narrowly to minimize the

number of backups that need to be retained and eventually restored for discovery purposes.

.A preservation order may be difficult to implement perfectly and may cause hardship when the records are stored in data-processing systems that automatically control the period of retention. Revision of existing computer programs to provide for longer retention, even if possible, may be prohibitively expensive. Consider alternatives, such as having parties duplicate relevant data on removable media or retaining periodic backups.

Any preservation order should ordinarily permit destruction after reasonable notice to opposing counsel; if opposing counsel objects, the party seeking destruction should be required to show good cause before destruction is permitted. The order may also exclude specified categories of documents or data whose cost of preservation outweighs substantially their relevance in the litigation, particularly if copies of the documents or data are filed in a document depository * * * or if there are alternative sources for the information. The court can defer destruction if relevance cannot be fairly evaluated until the litigation progresses. As issues in the case are narrowed, the court may reduce the scope of the order. The same considerations apply to the alteration or destruction of physical evidence.” Manual, §11.442 (footnote omitted).

4. Address the consequences of inadvertent production of privileged materials.⁷
 5. Learn areas of agreement/disagreement about “subjects on which discovery may be needed.” Rule 26(f)(2).
 6. Plan your discovery requests.
- E. Thoughts from the Manual on what might be done by attorneys:

“The time and expense of discovery may sometimes be substantially reduced if pertinent information can be retrieved from existing computerized records. Moreover, production in computer-readable form of relevant files and fields (or even

of an entire database) can reduce disputes over the accuracy of compilations made from such data and enable experts for both sides to conduct studies using a common set of data. On the other hand, accessing and using computer-generated evidence is subject to numerous pitfalls. * * *. The parties' computer experts should informally discuss, in person or by telephone, procedures to facilitate retrieval and production of computerized information; the attorneys can then confirm these arrangements in writing." Manual, §32.432 (footnote omitted).

IV. DIGITAL INFORMATION AND RULE 26(a)(1)

A. Rule 26(a)(1) requires the automatic disclosure of, among other things:

“(A) the name and if known, the address and telephone number of each individual likely to have discoverable information that the disclosing party may use to support its claims or defenses, unless solely for impeachment, identifying the subjects of the information;

(B) a copy of, or a description by category and location of, all documents, data compilations, and tangible things that are in the possession, custody, or control of the party and that the disclosing party may use to support its claims or defenses, unless solely for impeachment;

(C) a computation of any category of damages claimed by the disclosing party, making available for inspection and copying as under Rule 34 the documents or other evidentiary material, not privileged or protected from disclosure, on which such computation is based, including materials bearing on the nature and extent of injuries suffered ***.” (emphasis added).

B. Rule 26(a)(1) also introduces the concept of bifurcation of discovery. Rule 26(a)(1) requires disclosure information that a party “may use to support its claims or defenses. “ This is consistent with Rule 26(b)(1), which allows discovery “regarding any matter *** that is relevant to the claim or defense of any party.” Attorneys should use the Rule 26(f) meeting to decide what the “claims or defenses” in a case are and the nature of Rule 26(a)(1) disclosure of electronic information.

C. Is the individual most knowledgeable about a party’s electronic information systems subject to disclosure under Rule 26(a)(1)(A)?

D. Rule 26(a) allows a party to object to disclosure:

“These disclosures must be made at or within 14 days after the Rule 26(f) conference unless a different time is set by stipulation or court order, or unless a party objects during the conference that initial disclosures are not appropriate in the circumstances of the action and states the objection in the

Rule 26(f) discovery plan. In ruling on the objection, the court must determine what disclosures—if any—are to be made, and set the time for disclosure.”

E. What the Manual says:

“Prediscovery disclosure avoids the cost of unnecessary formal discovery and accelerates the exchange of basic information to plan and conduct discovery and settlement negotiations. The judge should administer Rule 26(a)(1) to serve those purposes; disclosure should not place unreasonable or unnecessary burdens on the parties (and should not require disclosure of any information that would not have to be disclosed in response to formal discovery requests). In complex litigation, this rule may need modification or suspension.

The scope of disputed issues and relevant facts in a complex case may not be sufficiently clear from the pleadings to enable parties to make the requisite disclosure. One purpose of Rule 26(f)’s required meeting of counsel is to identify issues and reach agreement on the content and timing of the initial disclosures. To the extent the parties cannot agree at their meeting, it sometimes helps to defer disclosure and fashion an order at the Rule 16 conference, defining and narrowing the factual and legal issues in dispute and establishing the scope of disclosure. This will require suspending, by stipulation or order, Rule 26(f)’s presumptive ten-day deadline for making disclosure.

Although Rule 26(a)(1) defines certain information that must be disclosed, it does not limit the scope of prediscovery disclosure and exchange of information. The parties have a duty to conduct a reasonable investigation pursuant to disclosure, particularly when a party possesses extensive computerized data, which may be subject to disclosure or later discovery. The rule does not require actual production (except for damage computations and insurance agreements), but only identification of relevant information and materials. The judge may nevertheless direct the parties to produce and exchange materials in advance of discovery,

subject to appropriate objections. Effective use of this device without excessive and unnecessary burdens on the parties can streamline the litigation.” Manual, §11.13 (footnote omitted).

V. DISCOVERY

A. Basics

1. Back to the bifurcation of discovery: In addition to allowing discovery on any matter “relevant to the claim or defense,” Rule 26(b)(1) allows discovery, “[f]or good cause *** of any matter relevant to the subject matter involved in the action.” The bifurcation was introduced by the 2000 amendment of Rule 26(b)(1) and, according to the Advisory Committee Note, “is designed to involve the court more actively in regulating the breadth of sweeping or contentious discovery.” Unfortunately, as the Advisory Committee Note goes on to say, “[t]he dividing line between information relevant to the claims and defenses and that relevant only to the subject matter of the action cannot be defined with precision.”

2. The concept of proportionality. This appears in Rule 26(b)(2), which provides, in pertinent part:

“The frequency or extent of use of the discovery methods otherwise permitted under these rules and by any local rule shall be limited by the court if it determines that: (I) the discovery sought is unreasonably cumulative or duplicative, or is obtainable from some other source that is more convenient, less burdensome, or less expensive; (ii) the party seeking discovery has had ample opportunity by discovery in the action to obtain the information sought; or (iii) the burden or expense of the proposed discovery outweighs its likely benefit, taking into account the needs of the case, the amount in controversy, the parties’ resources, the importance of the issues at stake in the litigation, and the importance of the proposed discovery in resolving the issues. The court may act upon its own initiative after reasonable notice or pursuant to a motion under Rule 26(c).”

Rule 26(b)(2) “contemplates greater judicial involvement in the discovery process and thus acknowledges the reality that it cannot always operate on a self-regulating basis.” Advisory Committee Note to 1983 amendments to Rule 26. “The objective is to guard against redundant or disproportionate discovery ***.” *Id.* By 2000, the Advisory Committee “has been told repeatedly that courts have not implemented these limitations with the vigor that was contemplated.” GAP Report of Advisory Committee to 2000 amendment to Rule 26(b)(1). 192 F.R.D. 340, 390 (2000).⁸

3. Rule 26(g) (“Signing of Disclosures, Discovery Requests, Responses, and Objections”):

“(1) Every disclosure made pursuant to subdivision (a)(1) or subdivision (a)(3) shall be signed by at least one attorney of record in the attorney’s individual name, whose address shall be stated. An unrepresented party shall sign the disclosure and state the party’s address. The signature of the attorney or party constitutes a certification that to the best of the signer’s knowledge, information, and belief, formed after a reasonable inquiry, the disclosure is complete and correct as of the time it is made.

(2) Every discovery request, response, or objection made by a party represented by an attorney shall be signed by at least one attorney of record in the attorney’s individual name, whose address shall be stated. An unrepresented party shall sign the request, response, or objection and state the party’s address. The signature of the attorney or party constitutes a certification that to the best of the signer’s knowledge, information, and belief, formed after a reasonable inquiry, the request, response, or objection is:

(A) consistent with these rules and warranted by existing law or a good faith argument for the extension, modification, or reversal of existing law;

(B) not interposed for any improper purpose, such as to harass or to cause unnecessary delay or needless increase in the cost of litigation; and

(C) not unreasonable or unduly burdensome or expensive, given the needs of the case, the discovery already had in the case, the amount in controversy, and the importance of the issues at stake in the litigation.

If a request, response, or objection is not signed, it shall be stricken unless it is signed promptly after the omission is called to the attention of the party making the request, response, or objection, and a party shall not be obligated to

take any action with respect to it until it is signed. ***.”

4. Rule 34 (“Production of Documents and Things”):

“(a) Scope. Any party may serve on any other party a request (1) to produce and permit the party making the request, or someone acting on the requestor’s behalf, to inspect and copy, any designated documents (including writings, drawings, graphs, charts, photographs, phone records, and other data compilations from which information can be obtained, translated, if necessary, by the respondent through detection devices into reasonably usable form), or to inspect and copy, test, or sample any tangible things which constitute or contain matters within the scope of Rule 26(b) and which are in the possession, custody or control of the party upon whom the request is served; or (2) to permit entry upon designated land or other property in the possession or control of the party upon whom the request is served for the purpose of inspection and measuring, surveying, photographing, testing, or sampling the property or any designated object or operation thereon, within the scope of Rule 26(b).”

What the Advisory Committee Note to the 1970 amendment to Rule 34 says about “documents:”

“The inclusive description of ‘documents’ is revised to accord with changing technology. It makes clear that Rule 34 applies to electronics data compilation from which information can be obtained only with the use of detection devices, and that when the data can as a practical matter be made usable by the discovering party only through respondent’s devices, respondent may be required to use his devices to translate the data into usable form. In many instances, this means that respondent will have to supply a print-out of computer data. The burden thus placed on respondent will vary from case to case, and the courts have ample power under Rule 26(c) to protect respondent against undue burden or expense, either by restricting discovery or requiring that the discovering party pay costs. Similarly, if the discovering party needs to check the electronic source

itself, the court may protect respondent with respect to preservation of his records, confidentiality of nondiscoverable matters, and costs.”

In re: Ford Motor Co., 345 F.3d 1315 (11th Cir. 2003):

“*** Rule 34(a) does not grant unrestricted, direct access to a respondent’s database compilations. Instead, Rule 34(a) allows a requesting party to inspect and to copy the product—whether it be a document, disk, or other device—resulting from the respondent’s translation of the data into a reasonably useful form.

*** Like the other discovery rules, Rule 34(a) allows the responding party to search his records to produce the required, relevant data. Rule 34(a) does not give the requesting party the right to conduct the actual search. While at times—perhaps due to improper conduct on the part of the responding party—the requesting party itself may need to check the data compilation, the district court must ‘protect respondent with respect to the preservation of his records, confidentiality of nondiscoverable matters, and costs.’” 345 F.3d at 1316-17. (quoting Rule 34(a)).⁹

What the Manual says about production of computerized data under Rule 34:

“Conventional ‘warehouse’ productions of paper documents often were costly and time consuming, but the burdens and expense were kept in check by the time and resources available to the requesting parties to review and photocopy the documents. In a computerized environment, the relative burdens and expense shift dramatically to the responding party. The cost of searching and copying electronic data is insignificant. Meanwhile, the tremendously increased volume of computer data and a lack of fully developed electronic records-management procedures have driven up the cost of locating, organizing, and screening data for relevance and privilege prior to production. Allowing requesting parties access to the responding parties’ computer systems to conduct their own searches, which is in one sense analogous to the conventional warehouse paper production,

would compromise legally recognized privileges, trade secrets, and often the personal privacy of employees and customers.” Manual, §11.446.

B. Cost-bearing.

1. In 1998, the Advisory Committee proposed an amendment to Rule 34(b). The amendment would have added this sentence:

“On motion under Rule 37(a) or Rule 26(c) , or on its own motion, the court shall-if appropriate to implement the limitations of rule 26(b)(i)(iii)-limit the discovery or require the party seeking discovery to pay part or all of the reasonable expenses incurred by the responding part.” 181 F.R.D. 18, 88-89.

2. The intent of the Advisory Committee was to make “explicit the court’s authority to condition document production on payment by the party seeking discovery of part or all of the reasonable costs of that document production if the request exceeds the limitations of Rule 26(b)(2)(i), or (iii). This authority was implicit in the 1983 adoption of Rule 26(b)(2) ***.” 181 F.R.D. 18, 89-91 (1999).

3. The Judicial Conference did not approve the amendment. However, the power to shift costs remains implicit in Rules 26(b)(2) and 26(c). See Manual, §11.433; 8 Wright, Miller & Marcus, Federal Practice and Procedure, §2008.1 at 25 (2003 pocket part).

VI. COST-BEARING: THREE APPROACHES

A. McPeck v. Ashcroft, 202 F.R.D. 31 (D.D.C. 2001) (Magistrate Judge John M. Facciola).

1. Background of case and discovery dispute:

“Plaintiff’s complaint identifies two forms of retaliation. He first complains that, despite the confidentiality of the settlement agreement, his claims *** were known by the people with whom he worked and that he suffered humiliation and retaliation at their hands. He then complains that, after hiring counsel in July 1988 to pursue formal legal remedies beginning with EEO counseling, he suffered renewed retaliation efforts. ***.

In responding to plaintiff’s discovery, defendants have searched for electronic and paper documents. Since defendants have already searched for electronic records, they do not quarrel with their obligation to do so. During discovery, the producing party has an obligation to search available electronic systems for the information demanded. * * * Plaintiff, however, wants more. He wants to force DOJ to search its backup systems since they might yield, for example, data that was ultimately deleted by the user but was stored on the backup tape and remains there today.

Defendants protest that the remote possibility that such a search will yield relevant evidence cannot possibly justify the costs involved”. 202 F.R.D. at 32.

2. Judge Facciola’s analysis of cost-bearing:

“There is certainly no controlling authority for the proposition that restoring all backup tapes is necessary in every case. The Federal Rules of Civil Procedure do not require such a search, and the handful of cases are idiosyncratic and provide little guidance. The one judicial rationale that has emerged is that producing backup tapes is a cost of doing business in the computer age. ***. But, that assumes an alternative. It is impossible to walk ten feet into

the office of a private business or government agency without seeing a network computer, which is on a server, which, in turn, is being backed up on tape (or some other media) on a daily, weekly or monthly basis. What alternative is there? Quill pens?

Furthermore, making the producing party pay for all costs of restoration as a cost of its 'choice' to use computers creates a disincentive for the requesting party to demand anything less than all of the tapes. American lawyers engaged in discovery have never been accused of asking for too little. To the contrary, like the Rolling Stones, they hope that if they ask for what they want, they will get what they need. They hardly need any more encouragement to demand as much as they can from their opponent.

The converse solution is to make the party seeking the restoration of the backup tapes pay for them, so that the requesting party literally gets what it pays for. Those who favor a 'market' economic approach to the law would argue that charging the requesting party would guarantee that the requesting party would only demand what it needs. Under that rationale, shifting the cost of production solves the problem. ***.

But there are two problems with that analysis. First, a strict cost-based approach ignores the fact that a government agency is not a profit-producing entity and it cannot be said that paying costs in this case would yield the same 'profit' that other foregone economic activity would yield. ***. While the notion that government agencies and businesses will not have backup systems if they are forced to restore them whenever they are sued may seem fanciful, courts should not lead them into temptation.

Second, if it is reasonably certain that the backup tapes contain information that is relevant to a claim or defense, shifting all costs to the requesting party means that the requesting party will have to pay for the agency to search the backup tapes even though the requesting party would not have to pay for such a search of a 'paper' depository.

A fairer approach borrows, by analogy, from the economic principle of ‘marginal utility’. The more likely it is that the backup tape contains information that is relevant to a claim or defense, the fairer it is that the government agency search at its own expense. The less likely it is, the more unjust it would be to make the agency search at its own expense. The difference is ‘at the margin.’

Finally, economic considerations have to be pertinent if the court is to remain faithful to its responsibility to prevent “undue burden or expense”. Fed. R. Civ. P. 26(c). If the likelihood of finding something was the only criterion, there is a risk that someone will have to spend hundreds of thousands of dollars to produce a single e-mail. That is an awfully expensive needle to justify searching a haystack. It must be recalled that ordering the producing party to restore backup tapes upon a showing of likelihood that they will contain relevant information in every case gives the plaintiff a gigantic club with which to beat his opponent into settlement. No corporate president in her right mind would fail to settle a lawsuit for \$100,000 if the restoration of backup tapes would cost \$300,000. While that scenario might warm the cockles of certain lawyers’s hearts, no one would accuse it of being just.

Given the complicated questions presented, the clash of policies and the lack of precedential guidance, I have decided to take small steps and perform, as it were, a test run. Accordingly, I will order DOJ to perform a backup restoration of the e-mails attributable to *** [an individual’s] computer during the period of July 1, 1988 to July 1, 1999. I have chosen this period because a letter from plaintiff’s counsel to DOJ, complaining of retaliation and threatening to file an administrative claim, is dated July 2, 1998, and it seems to me a convenient and rational starting point to search for evidence of retaliation. I have chosen e-mail because of its universal use and because I am hoping that the restoration will yield both the e-mails *** [the individual] sent and those he received.

The DOJ will have to carefully document the time and money spent in doing the search. It will then have to search in the restored e-mails for any document responsive to any of plaintiff's requests for production of documents. Upon the completion of this search, the DOJ will then file a comprehensive, sworn certification of the time and money spent and the results of the search. Once it does, I will permit the parties an opportunity to argue why the results and the expense do or do not justify any further search." 202 F.R.D. at 33-35 (citations omitted).¹⁰

B. Rowe Entertainment, Inc. v. William Morris Agency, Inc., 205 F.R.D. 421 (S.D.N.Y.), aff'd, 53 Fed. R. Serv. 3d 296 (S.D.N.Y. 2002) (Magistrate Judge James C. Francis).

1. Background of case and discovery dispute:

"Too often, discovery is not just about uncovering the truth, but also about how much of the truth the parties can afford to disinter. As this case illustrates, discovery expenses frequently escalate when information is stored in electronic form.

The plaintiffs are black concert promoters who contend that they have been frozen out of the market for promoting events with white bands by the discriminatory and anti-competitive practices of the defendants." 205 F.R.D. at 423.

"The plaintiffs' document demands are sweeping. For example, they demand production of all documents concerning any communication between any defendants relating to the selection of concert promoters and bids to promote concerts. *** Similarly, the plaintiffs have requested '[a]ll documents concerning the selection of concert promoters, and the solicitation, and bidding processes relating to concert promotions.' *** They have also demanded '[a]ll documents concerning market shares, market share values, market conditions, or geographic boundaries in which any ... concert promoter operates.' These are but three examples of

the thirty-five requests made in the plaintiffs' first document demand."

Each of the moving defendants contends that it should be relieved of the obligation of producing e-mail responsive to the plaintiffs' requests because the burden and expense involved would far outweigh any possible benefit in terms of discovery of additional information. If production is nevertheless required, the defendants ask that the plaintiffs bear the cost. ***. 205 F.R.D. at 424.

2. Was the information sought discoverable?

"The plaintiffs have successfully demonstrated that the discovery they seek is generally relevant. Although the defendants vigorously contest the plaintiffs' interpretation of the documents that have already been produced *** those documents are plainly pertinent to the plaintiffs' claims. To the extent that the defendants' e-mails contain similar information, they are equally discoverable. ***."

Nor are the defendants' claims that the e-mail is unlikely to yield relevant information persuasive. General representations *** that *** employees do little business by e-mail are undocumented and are contradicted by data proffered by these same defendants. ***.

Furthermore, the supposition that important e-mails have been printed in hard copy form is likewise unsupported. In general, nearly one-third of all electronically stored data is never printed out. ***. Here, the defendants have not alleged that they had any corporate policy defining which e-mail messages should be reduced to hard copy because they are 'important.' Finally, to the extent that any employee of the defendants was engaged in discriminatory or anti-competitive practices, it is less likely that communications about such activities would be memorialized in an easily accessible form such as a filed paper document.

The defendants' concern about privacy is also

unavailing. To the extent that the corporate defendants' own privacy interests are at issue, they are adequately protected by the confidentiality order in this case. To the degree the defendants seek to assert the privacy concerns of their employees, those interests are severely limited. Although personal communications of employees may be [sic] appear in hard copy as well as in electronic documents ***, the defendants made no effort to exclude personal messages from the search of paper records conducted by plaintiffs' counsel. Moreover, an employee who uses his or her employer's computer for personal communications assumes some risk that they will be accessed by the employer or by others.

Thus, there is no justification for a blanket order precluding discovery of the defendants' e-mails on the ground that such discovery is unlikely to provide relevant information or will invade the privacy of non-parties." 205 F.R.D. at 428 (citations omitted).

3. Judge Francis' analysis of cost-bearing:

"The more difficult issue is the extent to which each party should pay the costs of production. 'Under [the discovery] rules, the presumption is that the responding party must bear the expense of complying with discovery requests [.]' ***. Nevertheless, a court may protect the responding party from 'undue burden or expense' by shifting some or all of the costs of production to the requesting party. ***. Here, the expense of locating and extracting responsive e-mails is substantial, even if the more modest estimates of the plaintiffs are credited. Therefore, it is appropriate to determine which, if any, of these costs, are 'undue,' thus justifying allocation of those expenses to the plaintiffs.

One line of argument, adopted by the plaintiffs, holds that the responding party should bear the costs of producing electronic data since 'if a party chooses an electronic storage method, the necessity for a retrieval program or method is an ordinary and foreseeable risk.' ***. But even if this principle

is unassailable in the context of paper records, it does not translate well into the realm of electronic data. The underlying assumption is that the party retaining information does so because that information is useful to it, as demonstrated by the fact that it is willing to bear the costs of retention. That party may therefore be expected to locate specific data, whether for its own needs or in response to a discovery request. With electronic media, however, the syllogism breaks down because the costs of storage are virtually nil. Information is retained not because it is expected to be used, but because there is no compelling reason to discard it. And, even if data is retained for limited purposes, it is not necessarily amenable to discovery. ***. Thus, it is not enough to say that because a party retained electronic information, it should necessarily bear the cost of producing it.

The contrary argument is that the requesting party should bear the burden since, when the costs of discovery are internalized, that party can perform a cost-benefit analysis and decide whether the effort is justified. ***. Yet, this ‘market’ approach has two shortcomings. First, it flies in the face of the well-established legal principle, cited above, that the responding party will pay the expenses of production. Second, it places a price on justice that will not always be acceptable: it would result in the abandonment of meritorious claims by litigants too poor to pay for necessary discovery.

Because of the shortcomings of either bright-line rule, courts have adopted a balancing approach taking into consideration such factors as: (1) the specificity of the discovery requests; (2) the likelihood of discovering critical information; (3) the availability of such information from other sources; (4) the purposes for which the responding party maintains the requested data; (5) the relative benefit to the parties of obtaining the information; (6) the total cost associated with production; (7) the relative ability of each party to control costs and its incentive to do so; and (8) the resources available to each party. Each of these factors is relevant in determining whether discovery costs should be shifted in this case.” 205 F.R.D. at 428-29 (citations

omitted).

C. Zubulake v. UBS Warburg LLC, 217 F.R.D. 309 (S.D.N.Y. 2003) (District Judge Shira A. Scheindlin).

1. Background of case and discovery dispute:

“This case provides a textbook example of the difficulty of balancing the competing needs of broad discovery and manageable costs. Laura Zubulake is suing *** for gender discrimination and illegal retaliation. Zubulake’s case is certainly not frivolous and if she prevails, her damages may be substantial. She contends that key evidence is located in various e-mails exchanged among UBS employees that now exist only on backup tapes and perhaps other archived media. According to UBS, restoring those e-mails would cost approximately \$175,000.00, exclusive of attorney time in reviewing the e-mails. Zubulake now moves for an order compelling UBS to produce those e-mails at its expense.

At issue here is request number twenty-eight, for ‘all documents concerning any communication by or between UBS employees concerning Plaintiff.’ The term document in Zubulake’s request ‘includ[es], without limitation, electronic or computerized data compilations.’ ‘On July 8, 2002, UBS responded by producing approximately 350 pages of documents, including approximately 100 pages of e-mails. UBS also objected to a substantial portion of Zubulake’s requests.’”

“***. UBS, however, produced no additional e-mails and insisted that its initial production (the 100 pages of e-mails) was complete. As UBS’s opposition to the instant motion makes clear—although it remains unsaid—UBS never searched for responsive e-mails on any of its backup tapes. To the contrary, UBS informed Zubulake that the cost of producing e-mails on back-up tapes would be prohibitive (estimated at the time at approximately \$300,000.00).

“Zubulake, *** objected to UBS’s nonproduction. In fact, Zubulake knew that there were additional responsive e-mails that UBS had failed to produce because she herself had produced approximately 450 pages of e-mail correspondence. Clearly, numerous responsive e-mails had been created and deleted at UBS, and Zubulake wanted them.” 217 F.R.D. at 311-13 (footnotes omitted).

2. Was the information sought discoverable?

“*** Zubulake is entitled to discovery of the requested e-mails so long as they are relevant to her claims, which they clearly are. As noted, e-mail constituted a substantial means of communication among UBS employees. To that end, UBS has already produced approximately 100 pages of e-mails, the contents of which are unquestionably relevant.

Nonetheless, UBS argues that Zubulake is not entitled to any further discovery because it already produced all responsive documents, to wit, the 100 pages of e-mails. This argument is unpersuasive for two reasons. First, because of the way that UBS backs up its e-mail files, it clearly could not have searched all of its e-mails without restoring the ninety-four backup tapes (which UBS admits that it has not done). UBS therefore cannot represent that it has produced all responsive e-mails. Second, Zubulake herself has produced over 450 pages of relevant e-mails, including e-mails that would have been responsive to her discovery requests but were never produced by UBS. These two facts strongly suggest that there are e-mails that Zubulake has not received that reside on UBS’s backup media.” 217 F.R.D. at 317 (footnotes omitted).

3. “Should Cost-Shifting Be Considered?”

“Because it apparently recognizes that Zubulake is entitled to the requested discovery, UBS expends most of its efforts urging the court to shift the cost of production to ‘protect [it] ...from undue burden or expense.’ Faced with similar applications, courts generally in some sort of cost-

shifting analysis, whether the refined eight-factor Rowe test or a cruder application of Rule 34's proportionality test, or something in between.

The first question, however, is whether cost-shifting must be considered in every case involving the discovery of electronic data, which—in today's world— includes virtually all cases. In light of the accepted principle *** that electronic evidence is no less discoverable than paper evidence, the answer is, 'No.' The Supreme Court has instructed that 'the presumption is that the responding party must bear the expense of complying with discovery requests. ...' Any principled approach to electronic evidence must respect this presumption.

Courts must remember that cost-shifting may effectively end discovery, especially when private parties are engaged in litigation with large corporations. As large companies increasingly move to entirely paper-free environments, the frequent use of cost-shifting will have the effect of crippling discovery in discrimination and retaliation cases. This will both undermine the 'strong public policy favor[ing] resolving disputes on their merits,' and may ultimately deter the filing of potentially meritorious claims.

Thus, cost-shifting should be considered only when electronic discovery imposes an 'undue burden or expense' on the responding party. The burden or expense of discovery is, in turn, 'undue' when it 'outweighs its likely benefit, taking into account the needs of the case, the amount in controversy, the parties' resources, the importance of the issues at stake in the litigation, and the importance of the proposed discovery in resolving the issues.'

Many courts have automatically assumed that an undue burden or expense may arise simply because electronic evidence is involved. This makes no sense. Electronic evidence is frequently cheaper and easier to produce than paper evidence because it can be search automatically, key words can be run for privilege checks, and the production can be made in electronic form obviating the need for mass

photocopying.

In fact, whether production of documents is unduly burdensome or expensive turns primarily on whether it is kept in an accessible or inaccessible format (a distinction that corresponds closely to the expense of production). In the world of paper documents, for example, a document is accessible if it is readily available in a usable format and reasonably indexed. Examples of inaccessible paper documents could include (a) documents in storage in a difficult to reach place; (b) documents converted to microfiche and not easily readable; or (c) documents kept haphazardly, with no indexing system, in quantities that make page-by-page searches impracticable. But in the world of electronic data, thanks to search engines, any data that is retained in a machine readable format is typically accessible.” 217 F.R.D. at 317-18 (footnotes omitted).

4. Judge Scheindlin’s criticism of Rowe:

“In the year since Rowe was decided, its eight factor test has unquestionably become the gold standard for courts resolving electronic discovery disputes. But there is little doubt that Rowe factors will generally favor cost-shifting. Indeed, of the handful of reported opinions that apply Rowe or some modification thereof, all of them have ordered the cost of discovery to be shifted to the requesting party.

In order to maintain the presumption that the responding party pays, the cost-shifting analysis must be neutral; close calls should be resolved in favor of the presumption. The Rowe factors, as applied, undercut that presumption for three reasons. First, the Rowe test is incomplete. Second, courts have given equal weight to all of the factors, when certain factors should predominate. Third, courts applying the Rowe test have not always developed a full factual record” 217 F.R.D. at 320 (footnotes omitted).

“Certain factors specifically identified in the Rules are

omitted from Rowe's eight factors. In particular, Rule 26 requires consideration of 'the amount in controversy, the parties resources, the importance of the issues at stake in the litigation, and the importance of the proposed discovery in resolving the issues.' Yet Rowe makes no mention of either the amount in controversy or the importance of the issues at stake in the litigation. These factors should be added. Doing so would balance the Rowe factor that typically weighs most heavily in favor of cost-shifting, 'the total cost associated with production'. The cost of production is almost always an objectively large number in cases where litigating cost-shifting is worthwhile. But the cost of production when compared to 'the amount in controversy' may tell a different story. A response to a discovery request costing \$100,000 sound (and is) costly, but in a case potentially worth millions of dollars, the cost of responding may not be unduly burdensome.

Rowe also contemplates 'the resources available to each party.' But here too – although this consideration may be implicit in the Rowe test-the absolute wealth of the parties is not the relevant factor. More important than comparing the relative ability of a party to pay for discovery, the focus should be on the total cost of production as compared to the resources available to each party. Thus, discovery that would be too expensive for one defendant to bear would be a drop in the bucket for another.

Last, 'the importance of the issues at stake in the litigation' is a critical consideration, even if it is one that will rarely be invoked. For example, if a case has the potential for broad public impact, then public policy weighs heavily in favor of permitting extensive discovery. Cases of this ilk might include toxic tort class actions, environmental actions, so-called 'impact' or social reform litigation, cases involving criminal conduct, or cases implicating important legal or constitutional questions."

* * *

"Two of the Rowe factors should be eliminated.

First, the Rowe test includes ‘the specificity of the discovery request.’ Specificity is surely the touchstone of any good discovery request, requiring a party to frame a request broadly enough to obtain relevant evidence, yet narrowly enough to control costs. But relevance and cost are already two of the Rowe factors (the second and sixth). Because the first and second factors are duplicative, they can be combined. Thus, the first factor should be: the extent to which the request is specifically tailored to discover relevant information.

Second, the fourth factor, ‘the purposes for which the responding party maintains the requested data’ is typically unimportant. Whether the data is kept for a business purpose or for disaster recovery does not affect its accessibility, which is the practical basis for calculating the cost of production. Although a business purpose will often coincide with accessibility—data that is inaccessible is unlikely to be used or needed in the ordinary course of business—the concepts are not coterminous. In particular, a good deal of accessible data may be retained, though not in the ordinary course of business. For example, data that should rightly have been erased pursuant to a document retention/destruction policy may be inadvertently retained. If so, the fact that it should have been erased in no way shields that data from discovery. As long as the data is accessible, it must be produced.

Of course, there will be certain limited instances where the very purpose of maintaining the data will be to produce it to the opposing party. That would be the case, for example where the SEC requested ‘communications sent by [a] broker or dealer (including inter-office memoranda and communications) relating to his business as such.’ Such communications must be maintained ***. But in such cases, cost-shifting would not be applicable in the first place; the relevant statute or rule would dictate the extent of discovery and the associated costs. Cost-shifting would also be inappropriate for another reason—namely, that the regulation itself requires that the data be kept ‘in an accessible place.’” 217 F.R.D. at 321-22 (footnotes omitted).

5. Judge Scheindlin's analysis of cost-bearing:

“Set forth below is a new seven-factor test based on the modifications to Rowe discussed in the preceding sections.

1. The extent to which the request is specifically tailored to discover relevant information;
2. The availability of such information from other sources;
3. The total cost of production, compared to the amount in controversy;
4. The total cost of production, compared to the resources available to each party;
5. The relative ability of each party to control costs and its incentive to do so;
6. The importance of the issues at stake in the litigation; and
7. The relative benefits to the parties of obtaining the information.

* * *

Whenever a court applies a multi-factor test, there is a temptation to treat the factors as a check-list, resolving the issue in favor of whichever column has the most checks. But ‘we do not just add up the factors.’ When evaluating cost-shifting, the central question must be, does the request impose an ‘undue burden or expense’ on the responding party? Put another way, ‘how important is the sought-after evidence in comparison to the cost of production? The seven-factor test articulated above provide some guidance in answering this question, but the test cannot be mechanically applied at the risk of losing sight of its purpose.

Weighing the factors in descending order of importance may solve the problem and avoid a mechanistic application of the test. The first two factors—comprising the

marginal utility test—are the most important. ***.

The second group of factors addresses cost issues: ‘How expensive will this production be?’ and ‘Who can handle that expense?’ These factors include: (3) the total cost of production compared to the amount in controversy, (4) the total cost of production compared to the resources available to each party and (5) the relative ability of each party to control costs and its incentive to do so. The third ‘group’—(6) the importance of the litigation itself—stands alone, and as noted earlier will only rarely come into play. But where it does, this factor has the potential to predominate over the others. Collectively, the first three groups correspond to the three explicit considerations of Rule 26(b)(2)(iii). Finally, the last factor—(7) the relative benefits of production as between the requesting and producing parties—is the least important because it is fair to presume that the response to a discovery request generally benefits the requesting party. But in the unusual case where production will also provide a tangible or strategic benefit to the responding party, that fact may weigh against shifting costs.” 217 F.R.D. at 322-24 (footnotes omitted).¹¹

VII. AVOIDING PROBLEMS: SOME SUGGESTIONS

A. What the Manual suggests judges and attorneys can do:

“The judge should encourage the parties to discuss the scope of proposed computer-based discovery early in the case, particularly any discovery of data beyond that available to the responding parties in the ordinary course of business. The requesting parties should identify the information they require as narrowly and precisely as possible, and the responding parties should be forthcoming and explicit in identifying what data are available from what sources, to allow formulation of a realistic computer-based discovery plan. Rule 26(b)(2)(iii) allows the court to limit or modify the extent of otherwise allowable discovery if the burdens outweigh the likely benefit—the rule should be used to discourage costly, speculative, duplicative, or unduly burdensome discovery of computer data and systems. Additionally, some computerized data may have been compiled in anticipation of or for use in the litigation and may therefore be entitled to protection as trial preparation materials.”

* * *

“Evolving procedures use document-management technologies to minimize cost and exposure and, with time, parties and technology will likely continue to become more and more sophisticated. The judge should encourage the parties to discuss the issues of production forms early in litigation, preferably prior to any production, to avoid the waste and duplication of producing the same data in different formats. The relatively inexpensive production of computer-readable images may suffice for the vast majority of requested data. Dynamic data may need to be produced in native format, or in a modified format in which the integrity of the data can be maintained while the data can be manipulated for analysis. If raw data are produced, appropriate applications, file structures, manuals, and other tools necessary for the proper translation and use of the data must be provided. Files

(such as E-mail) for which metadata is essential to the understanding of the primary data should be identified and produced in an appropriate format. There may even be rare instances in which paper printouts (hard copy) are appropriate. No one form of production will be appropriate for all types of data in all cases.

Consider how to minimize and allocate the costs of production. Narrowing the overall scope of electronic discovery is the most effective method of reducing costs. Early agreement between the parties regarding the forms of production will help eliminate waste and duplication. More expensive forms of production, such as production of word-processing files with all associated metadata or production of data in a specified nonstandard format, should be conditioned upon a showing of need or sharing of expenses.” Manual, §11.446 (footnote omitted).

B. What the Manual says can be done to save time and expense:

“Phased or sequenced discovery of computerized data. Sections 11.41 and 11.422 have discussed phasing discovery by issue. Computerized data, however, are often not accessible by date, author, addressee, or subject matter without costly review and indexing. Therefore, it may be appropriate for the court to phase or sequence discovery of computerized data by accessibility. At the outset, allowing discovery of relevant, nonprivileged data available to the respondent in the routine course of business is appropriate and should be treated as a conventional document request. If the requesting party requests more computerized data, consider additional sources in ascending order of cost and burden to the responding party, e.g., metadata or system data, archived data, backup data, and legacy data. The judge should encourage the parties to agree to phased discovery of computerized data as part of the discovery plan. But with or without a prior agreement, the judge may engage in benefit-and-burden analysis under Rule 26(b)(2)(iii) at each stage and enter an appropriate order under Rule 26(c), which may include cost sharing between the parties or cost shifting to the requesting party * * *.

Computerized data produced in agreed-on-formats.

Information subject to discovery increasingly exists in digital or computer-readable form. The judge should encourage counsel to produce requested data in formats and on media that reduce transport and conversion costs, maximize the ability of all parties to organize and analyze the data during pretrial preparation, and assure usability at trial. Wholesale conversion of computerized data to paper form for production, only to be reconverted into computerized data by the receiving party, is costly and wasteful. Particularly in multiparty cases, data production on CD-ROM or by Internet-based data transfer can increase efficiency. Section 11.444 discusses ‘virtual’ document depositories.

Sampling of computer data. Parties may have vast collections of computerized data, such as stored E-mail messages or backup files containing routine business information kept for disaster recovery purposes. Unlike collections of paper documents, these data are not normally organized for retrieval by date, author, addressee, or subject matter, and may be very costly and time-consuming to investigate thoroughly. Under such circumstances, judges have ordered that random samples of data storage media be restored and analyzed to determine if further discovery is warranted under the benefit versus burden considerations of Rule 26(b)(2)(iii).” Manual, §11.423 (footnotes omitted).

C. What a district court should not do:

“In this case, Ford [the defendant] and Russell [the plaintiff] dispute whether Ford properly responded to Russell’s earlier requests for production. Although Russell asserts that Ford has not been forthright in providing documents, Ford contends that it has produced all relevant information. The district court was in the best position to determine whether Ford had improperly dealt with the earlier discovery requests. But the district court made no findings—express or implied—that Ford had failed to comply properly

with discovery requests.

The district court also did not discuss its view of Ford's objections and provided no substantive explanation for the court's ruling. Ford objected to the search on the grounds that (1) Russell had established no discovery abuses by Ford, (2) Ford had already searched the database and produced all relevant, non-privileged materials, and (3) the discovery rules did not allow the court to grant Russell free access to the databases regardless of relevance, privilege, or confidentiality. When a party objects to a motion for discovery, a court should rule on the objections and ordinarily give at least some statement of its reasons. ***.

Furthermore, in its order, the district court granted Russell unlimited, direct access to Ford's databases. The district court established no protocols for the search. The court did not even designate search terms to restrict the search. Without constraints, the order grants Russell access to information that would not—and should not—otherwise be discoverable without Ford first having had an opportunity to object.

While some kind of direct access might be permissible in certain cases, this case has not been shown to be one of those cases. Russell is unentitled to this kind of discovery without—at the outset—a factual finding of some non-compliance with discovery rules by Ford. By granting the sweeping order in this case, especially without such a finding, the district court clearly abused its discretion.” In re: Ford Motor Co., *supra*, 345 F. 3d at 1317.

D. In Metropolitan Opera Ass'n v. Local 100, 212 F.R.D. 178 (S.D.N.Y. 2003), egregious misconduct in discovery by the defendant union local led to the entry of judgment as to liability against it as well as other sanctions.¹² Here are some suggestions about what the defendant attorneys could have done to avoid the sanctions:

“The Metropolitan Opera decision does set out what the union should have done, at a minimum, to properly discharge its discovery obligations. Essentially, the court avers that the union had a duty to ‘establish a coherent and

effective system to faithfully and effectively respond to discovery requests.’ According to the court’s discussion, elements of that plan should have included:

- . a reasonable procedure to distribute discovery requests to all employees and agents of the defendant potentially possessing responsive information, and to account for the collection and subsequent production of the information to plaintiffs;

- . a method for explaining to their client what types of information would be relevant and responsive to discovery requests;

- . an inquiry into the client’s document retention or filing systems, and implementation of a systematic procedure for document production or for retention of documents, including electronic documents; and

- . proper supervision of all elements of discovery that were to be carried out by non-legal personnel.” V. Llewellyn, “The Court’s Prescription,” Vol. 3. No.3, Digital Discovery & e-Evidence 4 (March, 2003).¹³

E. Use of digital information at trial.¹⁴

1. What the Manual says:

“In general, the Federal Rules of Evidence apply to computerized data as they do to other types of evidence. Computerized data, however, raise unique issues concerning accuracy and authenticity. Accuracy may be impaired by incomplete data entry, mistakes in output instructions, programming errors, damage and contamination of storage media, power outages, and equipment malfunctions. The integrity of data may also be compromised in the course of discovery by improper search and retrieval techniques, data conversion, or mishandling. The proponent of computerized evidence has the burden of laying a proper foundation by establishing its accuracy.

The judge should therefore consider the accuracy and reliability of computerized evidence, including any necessary discovery during pretrial proceedings so that challenges to the evidence are not made for the first time at trial. When the data are voluminous, verification and correction of all items may not be feasible. In such cases, verification may be made of a sample of the data. Instead of correcting the errors detected in the sample—which might lead to the erroneous representation that the compilation is free from error—evidence may be offered (or stipulations made), by way of extrapolation from the sample, of the effect of the observed errors on the entire compilation. Alternatively, it may be feasible to use statistical methods to determine the probability and range of error.” Manual. §11.446 (footnote omitted).

2. Something to consider: admissibility of a facsimile transmission:

“The [District] Court was correct that ordinarily a fax’s sender would authenticate the document by testifying to such foundational facts as that the fax machine automatically date-stamps transmissions, that it was in proper working order, that she did not tamper with it, etc. ***. In this case Khorozian [the defendant] exercised her Fifth Amendment right against self-incrimination and thus did not take the stand. However, Kono [a witness] could—and did—authenticate the fax under Federal Rule of Evidence 901(a) by testifying that she received the fax on the date indicated on the header. Authentication does not conclusively establish the genuineness of an item; it is a foundation that a jury may reject.

Moreover, neither the header nor the text of the fax was hearsay. As to the header, ‘[u]nder FRE 901(a), a statement is something uttered by ‘a person,’ so nothing ‘said’ by a machine . . . is hearsay.’ ***. The fax contents were not hearsay because Khorozian sought to introduce the fax for the fact that it contained the name Teixeira (and was sent on May 15), not for its truth. The fax is relevant, regardless of its truth, to rebut the Government’s contention that she and Queirolo fabricated the document after May 25 as part of a scheme to defraud the bank.” United States v. Khorozian,

333 F.3d 498, 506 (3d Cir. 2003) (citations omitted).¹⁵

VIII. CONCLUSION

A. Learn about the relevant technology. “[C]ounsel must be cognizant of not only electronic discovery but also the details so that they can communicate effectively with clients, vendors, other counsel, and the courts.” J. Redgrave & E. Bachmann, “Ripples on the Shores of Zubulake,” The Federal Lawyer 31 (Nov./Dec. 2003).

B. Learn about the client’s information systems.

C. Make early-and specific-requests for discovery of digital information.

D. Use data sampling: “In a growing e-evidence trend, courts are looking to data sampling protocols-searching a small number of hard drives, servers, backup tapes, etc.-to see if relevant evidence exists * * *.” W. Furnish & M. Lange, “Lessons Learned: Rowe, Murphy Oil, Zubulake and Beyond,” Vol. 3 No. 12, Digital Discovery & e-Evidence 3 (Dec. 2003).

E. The case law on discovery disputes is fact-specific. Make the most complete record.¹⁶